

Entanglement Detection Using Majorization Uncertainty Bounds

M. Hossein Partovi*

Department of Physics and Astronomy, California State University, Sacramento, California 95819-6041

(Dated: March 2, 2013)

Entanglement detection criteria are developed within the framework of the majorization formulation of uncertainty. The primary results are two theorems asserting linear and nonlinear separability criteria based on majorization relations, the violation of which would imply entanglement. Corollaries to these theorems yield infinite sets of scalar entanglement detection criteria based on quasi-entropic measures of disorder. Examples are analyzed to probe the efficacy of the derived criteria in detecting the entanglement of bipartite Werner states. Characteristics of the majorization relation as a comparator of disorder uniquely suited to information-theoretical applications are emphasized throughout.

PACS numbers: 03.65.Ud, 03.67.Mn, 03.65.Ca

I. INTRODUCTION

Quantum measurements in general have indeterminate outcomes, with the results commonly expressed as a vector of probabilities corresponding to the set of outcomes. In case of noncommuting observables, the joint indeterminacy of their measurement outcomes has an inviolable lower bound, in stark contrast to classical expectations. This was discovered by Heisenberg in his desire to advance the physical understanding of the newly discovered matrix mechanics by relating the unavoidable disturbances caused by the act of measurement to the fundamental commutation relations of quantum dynamics [1]. Heisenberg's arguments relied on the statistical spread of the measured values of the observables to quantify uncertainty. This gave rise to the variance formulation of the uncertainty principle which remains a powerful source of intuition on the structure and spectral properties of microscopic systems. With the prospect of quantum computing and the development of quantum information theory in recent decades, on the other hand, the need for a measure of uncertainty that can better capture its information theoretical aspects, especially in dealing with noncanonical observables, has inspired new formulations. Among these are the entropic measure developed in the eighties, and the majorization formulation proposed recently. Uncertainty relations resulting from these formulations have found application to quantum cryptography, information locking, and entanglement detection, in addition to providing uncertainty limits [2, 3].

In this paper we develop applications of the majorization formulation of uncertainty introduced in Ref. [3] to the problem of entanglement detection. Deciding whether a given quantum state is entangled is a central problem of quantum information theory and known to be computationally intractable in general [4]. As a result, computationally tractable necessary conditions for separability, which provide a partial solution to this prob-

lem, have been the subject of active research in recent years. Among these, the Peres-Horodecki positive partial transpose criterion actually provides necessary and sufficient separability conditions for $2 \otimes 2$ and $2 \otimes 3$ dimensional systems, and necessary conditions otherwise. Other notable results are the reduction and global versus local disorder criteria, both necessary conditions in general [5]. An observable that has non-negative expectation values for all separable states and negative ones for a subset of entangled states provides an operational method of entanglement detection and is known as an entanglement witness [6]. It has also long been known that uncertainty relations can serve a similar purpose by providing inequalities that must be satisfied by separable states and if violated signal entanglement [7, 8].

In this paper we extend the majorization formulation of uncertainty developed in Ref. [3] to the problem of entanglement detection. As discussed in that paper (hereafter referred to as Paper I) and in the following, majorization as a comparator of uncertainty is qualitatively different from, and stronger than, scalar measures of uncertainty. Consequently, the entanglement detection results developed here represent a qualitative strengthening of the existing variance and entropic results. This will be evident, among other things, by the fact that they yield, as corollaries, an infinite class of entanglement detectors based on scalar measures.

As mentioned above, results of quantum measurements are in general probability vectors deduced from counter statistics, and an information theoretical formulation of uncertainty is normally based on an order of uncertainty defined on such vectors (see Paper I). Thus a scalar measure of uncertainty is commonly a real, non-negative function defined on probability vectors whose value serves to define the uncertainty in question. For example, the Shannon entropy function is the measure of uncertainty for the standard entropic formulation of uncertainty [9]. By contrast, majorization provides a partial order of uncertainty on probability vectors that is in general more stringent, and fundamentally stronger, than a scalar measure [10]. Note that, unlike scalar measures, majorization does not assign a quantitative measure of uncertainty to

*Electronic address: hpartovi@csus.edu

probability vectors, and as a partial order may find a pair of vectors to be incomparable.

A characterization of majorization that clarifies the foregoing statements can be attained by considering the quasi-entropic set of measures. These were defined in Paper I as the set of concave, symmetric functions defined on probability vectors, and include the Shannon, Tsallis, and (a subfamily of) Rényi entropies as special cases [11]. It is important to realize that the uncertainty order determined by one member of the quasi-entropic set for a given pair of probability vectors may contradict that given by another. While additional considerations may justify the use of, e.g., Shannon entropy in preference to the others, the foregoing observation clearly indicates the relative nature of the uncertainty order given by a specific measure, and immediately raises the following question: are there pairs of vectors for which all quasi-entropic measures determine the same uncertainty order? The answer is yes, and the common determination of the quasi-entropic set in such cases defines the uncertainty order given by majorization. What about the cases where there are conflicting determinations by the members of the quasi-entropic set? The majorization relation defines such pairs as incomparable, whence the “partial” nature of the order defined by majorization. We may therefore consider the majorization order to be equivalent to the collective order determination of the entire quasi-entropic set (see §IV). This characterization of the majorization relation clearly shows its standing vis-à-vis the scalar measures of uncertainty. More practical definitions of the majorization relation will be considered in §II.

The rest of this paper is organized as follows. In §II we review elements of the majorization formulation of uncertainty needed for our work and establish the notation. In §III we present the central results of this paper on entanglement detection, and in §IV we derive entire classes of entanglement detectors for quasi-entropic measures as corollaries to the theorems of §III. Concluding remarks are presented in §V. Details of certain mathematical proofs are given in the Appendix.

II. MAJORIZATION FORMULATION OF UNCERTAINTY

This section presents a review of the basic elements of majorization theory and the formulation of the uncertainty principle based on it. It follows the treatment given in Paper I.

A. Majorization

The basic element of our formulation is the majorization protocol for comparing the degree of uncertainty, or disorder, among probability vectors, i.e., sequences of non-negative numbers summing to unity [10]. By def-

inition, λ^1 is no less uncertain than λ^2 if λ^1 equals a mixture of the permutations of λ^2 . Then λ^1 is said to be majorized by λ^2 and written $\lambda^1 \prec \lambda^2$. An equivalent definition that flushes out the details of the foregoing is based on the vector λ^\downarrow which is obtained from λ by arranging the components of the latter in a nonincreasing order. Then, $\lambda^1 \prec \lambda^2$ if $\sum_{i=1}^j \lambda_i^{1\downarrow} \leq \sum_{i=1}^j \lambda_i^{2\downarrow}$ for $j = 1, 2, \dots, d-1$, where d is the larger of the two dimensions and trailing zeros are added where needed. As stated earlier, the majorization relation is a *partial* order, i.e., not every two vectors are comparable under majorization. As suggested in §I, two vectors are found to be incomparable when the difference in their degrees of uncertainty does not rise to the level required by majorization. As evident from the second definition given above, majorization requires the satisfaction of $N-1$ inequalities if the number of non-zero components of the majorizing vector is N . This accounts for the strength of the majorization relation as compared to scalar measures of disorder. Indeed as alluded to in §I, for any quasi-entropic function $F(\lambda)$, $\lambda^1 \prec \lambda^2$ implies $F(\lambda^1) \geq F(\lambda^2)$, but not conversely. On the other hand, if for *every* quasi-entropic function $F(\lambda)$ we have $F(\lambda^1) \geq F(\lambda^2)$, then $\lambda^1 \prec \lambda^2$ [10].

To establish uncertainty bounds, we need to characterize the greatest lower bound, or *infimum*, and the least upper bound, or *supremum*, of a set of probability vectors [12]. The *infimum* is defined as the vector that is majorized by every element of the set and in turn majorizes any vector with that property. The *supremum* is similarly defined. We will briefly outline the construction of the *infimum* here and refer the reader to Paper I for further details.

Given a set of probability vectors $\{\lambda^a\}_{a=1}^N$, consider the vector μ^{inf} defined by

$$\mu_0^{inf} = 0, \quad \mu_j^{inf} = \min \left(\sum_{i=1}^j \lambda_i^{1\downarrow}, \sum_{i=1}^j \lambda_i^{2\downarrow}, \dots, \sum_{i=1}^j \lambda_i^{N\downarrow} \right), \quad 1 \leq j \leq d_{max}, \quad (1)$$

where d_{max} is the largest dimension found in the set. The infimum is then given by

$$\lambda_i^{inf} = [\inf(\lambda^1, \lambda^2, \dots, \lambda^N)]_i = \mu_i^{inf} - \mu_{i-1}^{inf}, \quad (2)$$

where $1 \leq i \leq d_{max}$. The construction of the supremum starts with μ^{sup} in parallel with Eq. (1), but with “max” replacing “min,” and may require further steps detailed in Paper I.

It is worth noting here that the infimum (supremum) of a pair of probability vectors will in general be more (less) disordered than either. Important special cases are (i) one of the two majorizes the other, in which case the latter is the infimum and the former the supremum of the two, and (ii) the two are equal, in which case either is both the infimum and the supremum. Furthermore, a useful qualitative rule is that the more “different” are two probability vectors, the further will the infimum or supremum of the two be from at least one of them. Finally, we note that while the infimum or supremum of a

set of probability vectors always exists, it need not be a member of the set.

B. Measurement and uncertainty

A generalized measurement may be defined by a set of positive operators $\{\hat{E}_\alpha\}$ called measurement elements and subject to the completeness condition $\sum_\alpha \hat{E}_\alpha = \hat{1}$. The probability that outcome α turns up in a measurement of the state $\hat{\rho}$ is given by the Born rule $P_\alpha(\hat{\rho}) = \text{tr}[\hat{E}_\alpha \hat{\rho}]$. A generalized measurement can always be considered to be the restriction of a more basic type, namely a *projective* measurement, performed on an enlarged system to the system under generalized measurement [13]. A projective measurement is usually associated with an observable of the system represented by a self-adjoint operator \hat{M} , and entails a partitioning of the spectrum of \hat{M} into a collection of subsets $\{b_\alpha^M\}$ called measurement *bins*. We call a projective measurement *maximal* if each bin consists of a single point of the spectrum of the measured observable.

Having assembled the necessary concepts, we can now characterize uncertainty by means of majorization relations in a natural manner. To start, we define the probability vector $\mathcal{P}^X(\hat{\rho})$ resulting from a measurement X on a state $\hat{\rho}$ to be *uncertain* if it is majorized by $\mathcal{I} = (1, 0, \dots, 0)$ but not equal to it. As such, $\mathcal{P}^X(\hat{\rho})$ is said to be strictly majorized by \mathcal{I} and written $\mathcal{P}^X(\hat{\rho}) \prec \mathcal{I}$. Similarly, given a pair of measurements X and Y on a state $\hat{\rho}$, we say $\mathcal{P}^X(\hat{\rho})$ is more uncertain, equivalently more disordered, than $\mathcal{P}^Y(\hat{\rho})$ if $\mathcal{P}^X(\hat{\rho}) \prec \mathcal{P}^Y(\hat{\rho})$. Further, we define the joint uncertainty of a pair of measurements X and Y to be the outer product $\mathcal{P}^X \otimes \mathcal{P}^Y$, i.e., $\mathcal{P}_{\alpha\beta}^{X \otimes Y} = \mathcal{P}_\alpha^X \mathcal{P}_\beta^Y$. Since $H(\mathcal{P}^X \otimes \mathcal{P}^Y) = H(\mathcal{P}^X) + H(\mathcal{P}^Y)$, where $H(\cdot)$ is the Shannon entropy function, this definition is seen to be consistent with its entropic counterpart. As stated earlier, $\mathcal{P}^X \prec \mathcal{P}^Y$ implies $H(\mathcal{P}^X) \geq H(\mathcal{P}^Y)$ but not conversely. These definitions naturally extend to an arbitrary number of states and measurements.

We are now in a position to state the majorization statement of the uncertainty principle established in Paper I:

“The joint results of a set of generalized measurements of a given state are no less uncertain than a probability vector that depends on the measurement set but not the state, and is itself uncertain unless the measurement elements have a common eigenstate.”

In symbols,

$$\mathcal{P}^X(\hat{\rho}) \otimes \mathcal{P}^Y(\hat{\rho}) \otimes \dots \otimes \mathcal{P}^Z(\hat{\rho}) \prec \mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} \prec \mathcal{I}, \quad (3)$$

where

$$\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \sup_{\hat{\rho}} [\mathcal{P}^X(\hat{\rho}) \otimes \mathcal{P}^Y(\hat{\rho}) \otimes \dots \otimes \mathcal{P}^Z(\hat{\rho})], \quad (4)$$

unless the measurement elements $\{\hat{E}^X, \hat{E}^Y, \dots, \hat{E}^Z\}$ have a common eigenstate in which case $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$.

Note that $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ is the majorization uncertainty bound for the measurement set considered.

C. majorization uncertainty bounds

As stated above, the uncertainty bound $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ given by Eq. (4) depends on the measurement set but not the state of the system. As the supremum of all possible measurement outcomes, it is the probability vector that sets the irreducible lower bound to uncertainty for the set. As such, it is the counterpart of the variance product or entropic lower bound in the traditional formulations of the uncertainty principle. Unlike the latter, however, $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ is in general not realizable or even approachable by any state of the system. Therefore, there is in general no such thing as a “minimum uncertainty state” within the majorization framework. This is a consequence of the fact, mentioned earlier, that the infimum or supremum of a set of vectors need not be a member of the set. An obvious special case is the trivial example of zero uncertainty for which $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$, signaling the existence of a common eigenstate for the measurement elements.

Intuitively, we expect that mixing states can only increase their uncertainty, as is known to be the case for scalar measures of uncertainty. In case of majorization, this expectation is manifested in the property that the uncertainty bound $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ is in general realized on the class of pure states. Indeed for two and three mutually unbiased observables on a two-dimensional Hilbert space considered, we found in Paper I that the required maxima for the components of μ^{sup} that serve to define the uncertainty bound are reached on pure states. More specifically, as outlined in §IIA above and detailed in §IIB, IVA and IVB of Paper I, the calculation of $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ involves a component-wise determination of μ^{sup} by a series of maximizations over all possible density matrices $\hat{\rho}$. The property in question then guarantees that the maximization process can be limited to density matrices representing pure states only. We will prove this assertion in general, as well as for the class of separable states, in the Appendix.

III. ENTANGLEMENT DETECTION

As discussed earlier, an entanglement detector can be effective in deciding whether a given density matrix is separable by providing a condition that is satisfied by all separable states and if violated signals entanglement. A large body of entanglement detection strategies have been developed in recent years which are primarily based on scalar conditions, including those based on variance and entropic type uncertainty relations, which can be found in Refs. [5–8]. Here we shall develop majorization conditions for entanglement detection based on the formulation of measurement uncertainty given in Paper

I and outlined in §II above. In particular, the entanglement condition given in Theorem 1 below is linear and susceptible to experimental implementation, so it can be formulated as an entanglement witness. Nonlinear detectors developed below, on the other hand, rely on majorization-based uncertainty relations. We will also introduce the notion of subsystem disorder and a sharpened version of the Nielsen-Kempe [14] separability condition as a nonlinear entanglement detector.

As majorization relations, our results in general entail multiple inequalities whose number will grow with the uncertainty levels involved. As discussed in §IIA, this is an important feature of the majorization relation as comparator of disorder, one that sets it apart from scalar conditions and provides for the refinement needed in comparing highly disordered vectors. As will be seen in §IV, this property has the consequence that each majorization condition yields a scalar condition for the entire quasi-entropic class of uncertainty measures.

A. Linear detectors

Our detection strategy is based on the intuitive expectation that the measurement uncertainty bound for the class of separable states of a multipartite system must be majorized by the corresponding bound for all states, and that this hierarchy can be exploited for entanglement detection. This is the majorization rendition of the strategy often used to derive separability conditions [7, 8]. We will first consider the case of one generalized measurement resulting in a linear detection condition.

Theorem 1. Let the results of the generalized measurement $(X, \{\hat{E}_\alpha^X\})$ on a multipartite state $\hat{\rho}^{ABC\dots F}$ of parties (A, B, C, \dots, F) defined on a finite-dimensional Hilbert space be bounded by

$$\mathcal{P}_{sup}^X = \sup_{\hat{\rho}^{ABC\dots F}} \mathcal{P}^X(\hat{\rho}^{ABC\dots F}), \quad (5)$$

and in the case of a separable state $\hat{\rho}_{sep}^{ABC\dots F}$ by

$$\mathcal{P}_{sep;sup}^X = \sup_{\hat{\rho}_{sep}^{ABC\dots F}} \mathcal{P}^X(\hat{\rho}_{sep}^{ABC\dots F}). \quad (6)$$

Then, (i) the suprema in the foregoing pair of equations may be taken over the class of pure and pure, product states, respectively, and (ii) given an arbitrary state $\hat{\sigma}^{ABC\dots F}$, the condition $\mathcal{P}^X(\hat{\sigma}^{ABC\dots F}) \not\prec \mathcal{P}_{sep;sup}^X$ implies that $\hat{\sigma}^{ABC\dots F}$ is entangled [15].

Part (i) of Theorem 1 is the majorization version of the familiar result that uncertainty bounds are realized on pure states and is proved in the Appendix [16]. Part (ii) is the statement that the bound in Eq. (6) is a necessary condition for separability and its violation signals the existence of entanglement. As noted above, Theorem 1 provides an operational method of entanglement detection and can be reformulated as a set of entanglement witnesses.

Clearly, the efficacy of Theorem 1 in detecting entanglement depends on the choice of the measurement X for

a given quantum state. This suggests looking for the optimum measurement, in parallel with the optimization problem for entanglement witnesses [17]. Note that here the choice of the generalized measurement X affects both $\mathcal{P}^X(\hat{\sigma}^{ABC\dots F})$ and $\mathcal{P}_{sep;sup}^X$, thus making the task of finding a fully optimized solution rather difficult. It is therefore fortunate that we can achieve a partial optimization on the basis of Theorem 3 of Paper I, the relevant content of which we can state as follows:

Lemma Under the conditions of Theorem 1 and with X restricted to rank 1 measurements, we have

$$\mathcal{P}^X(\hat{\sigma}^{ABC\dots F}) \prec \mathcal{P}^{X^*}(\hat{\sigma}^{ABC\dots F}) = \lambda(\hat{\sigma}^{ABC\dots F}), \quad (7)$$

where X^* is a maximal projective measurement whose elements are the set of rank 1 orthogonal projection operators onto the eigenvectors of $\hat{\sigma}^{ABC\dots F}$ and $\lambda(\hat{\sigma}^{ABC\dots F})$ is the corresponding spectrum.

It should be noted here that while the choice of X^* succeeds in optimizing $\mathcal{P}^X(\hat{\sigma}^{ABC\dots F})$ which appears in part (ii) of Theorem 1, its effect on $\mathcal{P}_{sep;sup}^X$, the other object appearing therein, is unknown and may result in a poor overall choice. Moreover, any degeneracy in the spectrum of $\hat{\sigma}^{ABC\dots F}$ renders X^* nonunique and subject to further optimization. These features will be at play in the example considered below where, the foregoing caveats notwithstanding, the resulting detector turns out to be fully optimal.

The example in question is the Werner state $\hat{\rho}_d^{wer}(q)$ for a bipartite system of two d -level parties A and B defined on a d^2 -dimensional Hilbert space [18, 19]

$$\hat{\rho}_d^{wer}(q) = \frac{1}{d^2}(1-q)\hat{\mathbb{1}} + q|\mathfrak{B}_1\rangle\langle\mathfrak{B}_1|, \quad (8)$$

where

$$|\mathfrak{B}_1\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |A, j\rangle \otimes |B, j\rangle. \quad (9)$$

Here $\{|A, j\rangle\}$ and $\{|B, j\rangle\}$ are orthonormal bases for subsystem A and B respectively. Also, here and elsewhere we use the symbol $\hat{\mathbb{1}}$ to denote the identity operator of dimension appropriate to the context. Note that $|\mathfrak{B}_1\rangle$ is the generalization of the two-qubit Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$. As such, it is totally symmetric, as well as maximally entangled in the sense that its marginal states are maximally disordered and equal to $\hat{\mathbb{1}}/d$. An important fact to be used in the following is that $\hat{\rho}_d^{wer}(q)$ is separable if and only if $q \leq (1+d)^{-1}$ [19].

In order to probe the separability of $\hat{\rho}_d^{wer}(q)$ by means of Theorem 1, we must first choose a measurement. We will use the above Lemma to guide our choice of the putative optimum measurement X^* . The Lemma requires that the measurement elements be equal to the orthogonal projections onto the eigenvectors of $\hat{\rho}_d^{wer}(q)$. However, a simple calculation shows that the spectrum of $\hat{\rho}_d^{wer}(q)$ is degenerate, consisting of a simple eigenvalue equal to $q + d^{-2}(1-q)$ and a $(d^2 - 1)$ -fold degenerate eigenvalue equal to $d^{-2}(1-q)$. A natural choice of basis

for the degenerate subspace is a set of $d^2 - 1$ generalized Bell states which, together with $|\mathfrak{B}_1\rangle$ of Eq. (9), constitute the orthonormal eigenbasis $\{|\mathfrak{B}_\alpha\rangle\}_{\alpha=1}^{d^2}$ for $\hat{\rho}_d^{wer}(q)$. This generalized Bell basis is thus characterized by the fact that each $|\mathfrak{B}_\alpha\rangle$ is maximally entangled and possesses equal Schmidt coefficients.

The measurement elements of the optimal measurement X^* are thus given by $\hat{E}_\alpha^{X^*} = |\mathfrak{B}_\alpha\rangle\langle\mathfrak{B}_\alpha|$, $\alpha = 1, 2, \dots, d^2$. With X^* so defined, we readily find

$$\mathcal{P}^{X^*}[\hat{\rho}_d^{wer}(q)] = [q + d^{-2}(1-q), d^{-2}(1-q), d^{-2}(1-q), \dots, d^{-2}(1-q)], \quad (10)$$

which corresponds to the spectrum of $\hat{\rho}_d^{wer}(q)$ as expected. Note that the Lemma guarantees that any other rank 1 measurement in place of X^* would produce a more disordered probability vector on the right-hand side of Eq. (10).

The next step is the calculation of $\mathcal{P}_{sep;sup}^{X^*}$ of Theorem 1, which is the least disordered probability vector that can result from the measurement of X^* on a pure, product state of two d -level subsystems. As can be seen in Eq. (A.7) of the Appendix, this calculation requires finding the maximum overlap of every measurement element $|\mathfrak{B}_\alpha\rangle\langle\mathfrak{B}_\alpha|$ with the class of pure, product states. This overlap is given by the square of the largest Schmidt coefficient of the respective Bell state $|\mathfrak{B}_\alpha\rangle$ [20]. Since all Schmidt coefficients are equal to $1/\sqrt{d}$ for every Bell state $|\mathfrak{B}_\alpha\rangle$, we can conclude that

$$\mathcal{P}_{sep;sup}^{X^*} = (1/d, 1/d, \dots, 1/d, 0, 0, \dots, 0). \quad (11)$$

We are now in a position to apply the entanglement criterion of Theorem 1, which reads $\mathcal{P}^{X^*}[\hat{\rho}_d^{wer}(q)] \not\prec \mathcal{P}_{sep;sup}^{X^*}$ in this instance [21]. Using the information in Eqs. (10) and (11), we readily see that this criterion translates to the single inequality $q + d^{-2}(1-q) > 1/d$, or $q > 1/(1+d)$, which is the necessary and sufficient condition for the inseparability of the Werner state $\hat{\rho}_d^{wer}(q)$. Thus every entangled Werner state of two d -level systems is detected by the measurement X^* defined above.

While the efficacy of the generalized Bell states for detecting entanglement in Werner states is well established [5–8], we note its natural emergence from the general majorization results of this subsection. It should also be noted that Theorem 1 and the foregoing analysis can be extended to multipartite states.

B. Nonlinear detectors

As an example of nonlinear entanglement detectors, we will derive majorization conditions for separability based on uncertainty relations. The method we will follow relies on the degeneracy properties of projective measurements on a single system versus the products of such measurements on a multipartite system [7, 8]. For example, with measurement X^A on $\hat{\rho}^A$ and X^B on $\hat{\rho}^B$

and each having a simple spectrum, the product projective measurement $X^A \otimes X^B$ on the bipartite system $\hat{\rho}^{AB}$ may be degenerate. A specific case is $\hat{\sigma}_x^A \otimes \hat{\sigma}_x^B$, whose spectrum $(+1/4, +1/4, -1/4, -1/4)$ is doubly degenerate, versus $\hat{\sigma}_x^A$ or $\hat{\sigma}_x^B$ each of which has the simple spectrum $(+1/2, -1/2)$. Note the fact that here we are following common practice by defining $\hat{\sigma}_x^A \otimes \hat{\sigma}_x^B$ to be the projective measurement whose two elements project into the subspaces corresponding to the eigenvalues $+1/4$ and $-1/4$ (and not the rank 1 projective measurement that resolves the degeneracies by virtue of using all four product elements). Consequently, it may happen that $X^A \otimes X^B$ and $Y^A \otimes Y^B$ have a common eigenstate while X^A and Y^A do not, and that the said common eigenstate is an entangled pure state. In such cases, the corresponding probability vectors will reflect the stated differences and may be capable of detecting entanglement as in the case of linear detectors.

Consider two product projective measurements $X^A \otimes X^B$ and $Y^A \otimes Y^B$ performed on the bipartite state $\hat{\rho}^{AB}$. The measurement results are then given by

$$\begin{aligned} \mathcal{P}^{X^A \otimes X^B}(\hat{\rho}^{AB}) &= \text{tr}[\hat{\Pi}^{X^A \otimes X^B} \hat{\rho}^{AB}], \\ \mathcal{P}^{Y^A \otimes Y^B}(\hat{\rho}^{AB}) &= \text{tr}[\hat{\Pi}^{Y^A \otimes Y^B} \hat{\rho}^{AB}], \\ \mathcal{P}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}(\hat{\rho}^{AB}) &= \mathcal{P}^{X^A \otimes X^B}(\hat{\rho}^{AB}) \\ &\quad \otimes \mathcal{P}^{Y^A \otimes Y^B}(\hat{\rho}^{AB}), \end{aligned} \quad (12)$$

where $\hat{\Pi}^{X^A \otimes X^B}$ and $\hat{\Pi}^{Y^A \otimes Y^B}$, both projection operators but not necessarily rank 1, represent the measurement elements of the two product measurements. Note the majorization definition of joint uncertainty, given earlier, at work on the last line of Eq. (12). Under these conditions, we have the following general result.

Theorem 2. The results of measurements $X^A \otimes X^B$ and $Y^A \otimes Y^B$ on a separable state $\hat{\rho}_{sep}^{AB}$ satisfy

$$\mathcal{P}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}(\hat{\rho}_{sep}^{AB}) \prec \mathcal{P}_{sup}^{X \oplus Y}, \quad (13)$$

where $\mathcal{P}_{sup}^{X \oplus Y}$ is defined in Eq. (4), hence the condition $\mathcal{P}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}(\hat{\rho}^{AB}) \not\prec \mathcal{P}_{sup}^{X \oplus Y}$ implies that the state $\hat{\rho}^{AB}$ is entangled.

To establish Theorem 2, consider product states of the form $\hat{\rho}^A \otimes \hat{\rho}^B$. Then Lemma 1 of Ref. [8] asserts that

$$\begin{aligned} \mathcal{P}^{X^A \otimes X^B}(\hat{\rho}^A \otimes \hat{\rho}^B) &\prec \mathcal{P}^{X^A}(\hat{\rho}^A), \\ \mathcal{P}^{Y^A \otimes Y^B}(\hat{\rho}^A \otimes \hat{\rho}^B) &\prec \mathcal{P}^{Y^A}(\hat{\rho}^A). \end{aligned} \quad (14)$$

Using these relations and Eq. (12), we find [22]

$$\mathcal{P}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}(\hat{\rho}^A \otimes \hat{\rho}^B) \prec \mathcal{P}^{X^A}(\hat{\rho}^A) \otimes \mathcal{P}^{Y^A}(\hat{\rho}^A). \quad (15)$$

Since the right-hand side of Eq. (15) is by definition majorized by $\mathcal{P}_{sup}^{X \oplus Y}$, we conclude that

$$\mathcal{P}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}(\hat{\rho}^A \otimes \hat{\rho}^B) \prec \mathcal{P}_{sup}^{X \oplus Y}. \quad (16)$$

At this point we note that $\mathcal{P}_{sep;sup}^{(X^A \otimes X^B) \oplus (Y^A \otimes Y^B)}$, which by definition majorizes all probability vectors that can appear on the left-hand side of Eq. (13), can be found among pure, product states [16]. This assertion is established in Eq. (A.8) et seq. of the Appendix. Consequently, the product state $\hat{\rho}^A \otimes \hat{\rho}^B$ in Eq. (16) may be replaced by any separable state $\hat{\rho}_{sep}^{AB}$, thereby establishing Eq. (13) and Theorem 2.

We note in passing that, because the measurements considered in Theorem 2 are in general not rank 1, the Lemma which we used earlier for detector optimization cannot be applied here.

To illustrate Theorem 2, we will consider the case of three mutually unbiased observables measured on bipartite states of two-level systems, as in §IVB of Paper I. In effect, this amounts to measuring products of the three spin components of a pair of spin-1/2 systems (or qubits). For this case, Eq. (16) reads

$$\mathcal{P}^{(\sigma_x^A \otimes \sigma_x^B) \oplus (\sigma_y^A \otimes \sigma_y^B) \oplus (\sigma_z^A \otimes \sigma_z^B)}(\hat{\rho}_{sep}^{AB}) \prec \mathcal{P}_{sup}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}, \quad (17)$$

where $\hat{\rho}_{sep}^{AB}$ is any separable state of two qubits. Thus a violation of this relation by a two-qubit state implies that it is entangled.

For the set of states $\hat{\rho}^{AB}$ to be probed by Theorem 2, we will consider the Werner family of two-qubit states

$$\hat{\rho}^{wer}(q) = \frac{1}{4}(1-q)\mathbb{1} + q|\mathfrak{B}_1\rangle\langle\mathfrak{B}_1|, \quad (18)$$

which is known to be entangled if and only if $q > 1/3$. Here $0 \leq q \leq 1$ and $|\mathfrak{B}_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ is the totally symmetric Bell state of Eq. (9) for the present case.

A calculation of the probability vector for this measurement gives $(1 \pm q)(1 \pm q)(1 \pm q)/8$ for the 8 components of $\mathcal{P}^{(\sigma_x^A \otimes \sigma_x^B) \oplus (\sigma_y^A \otimes \sigma_y^B) \oplus (\sigma_z^A \otimes \sigma_z^B)}[\hat{\rho}^{wer}(q)]$. Arranged in a non-ascending order, these are $(1+q)^3/8$, $(1+q)^2(1-q)/8$ and $(1+q)(1-q)^2/8$ both three-fold degenerate, and $(1-q)^3/8$. The supremum $\mathcal{P}_{sup}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}$, on the other hand, was calculated in §IVB of Paper I, where it was found as

$$\begin{aligned} \mathcal{P}_{sup}^{\sigma_x \oplus \sigma_y \oplus \sigma_z} = \frac{1}{8} & [(1+1/\sqrt{3})^3, 2(1+1/\sqrt{2})^2 \\ & -(1+1/\sqrt{3})^3, 4 - (1+1/\sqrt{2})^2, 4 - (1+1/\sqrt{2})^2 \\ & , 0, 0, 0, 0]. \end{aligned} \quad (19)$$

According to Theorem 2, $\hat{\rho}^{wer}(q)$ violates the separability condition of Eq. (13) if $\mathcal{P}_{sup}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}$ fails to majorize $\mathcal{P}^{(\sigma_x^A \otimes \sigma_x^B) \oplus (\sigma_y^A \otimes \sigma_y^B) \oplus (\sigma_z^A \otimes \sigma_z^B)}[\hat{\rho}^{wer}(q)]$, which occurs for $q > 1/\sqrt{3} = 0.577$ in this example. The entanglement in $\hat{\rho}^{wer}(q)$ is thus detected by Theorem 2 for $q > 0.577$, and missed for the range $0.333 < q \leq 0.577$. By comparison, a similar method based on the Shannon entropy in Ref. [7] detects entanglement in $\hat{\rho}^{wer}(q)$ for $q > 0.65$. Reference [8], using the same method but relying on the family of Tsallis entropies, matches our

condition $q > 0.577$ by means of a numerical calculation that searches over the Tsallis family for optimum performance.

We end this subsection by a brief description of a sharpened version of the Nielsen-Kempe theorem as a nonlinear entanglement detector [12]. This celebrated theorem is an elegant separability condition based on majorization relations. It asserts that the spectrum of a separable bipartite state is majorized by each of its marginal spectra. By extension, this theorem guarantees that the spectra of all possible subsystems of a separable multipartite system must majorize its global spectrum [14]. Our version of this theorem employs the notion of the infimum of a set of probability vectors discussed in §IIA, and is based on the observation that if a probability vector is majorized by each member of a set of probability vectors, then it must also be majorized by the infimum of that set.

Consider the multipartite state $\hat{\rho}^{ABC\dots F}$ together with all its subsystem states $\{\hat{\rho}^{X_a}\}_{a=1}^f$ and subsystem spectra $\{\lambda^{X_a}\}_{a=1}^f$, where $\{X_a\}_{a=1}^f$ represent all proper subsets of the set of parties (A, B, C, \dots, F) . Then the infimum of the subsystem spectra, $\Lambda^{ABC\dots F} = \inf[\lambda^{X_1}, \lambda^{X_2}, \dots, \lambda^{X_f}]$, embodies the *subsystem disorder* of the state $\hat{\rho}^{ABC\dots F}$, as expressed in the following theorem.

Theorem 3 (Nielsen-Kempe). A multipartite state is entangled if its system disorder fails to exceed its subsystem disorder, i.e., $\lambda^{ABC\dots F}$ is entangled if $\lambda^{ABC\dots F} \not\prec \Lambda^{ABC\dots F}$.

IV. QUASI-ENTROPIC DETECTORS

The majorization based theorems of the previous section have direct corollaries that yield scalar entanglement detectors for the entire class of *Schur-concave* measures. A Schur-concave function G is defined by the property that $\lambda^1 \prec \lambda^2$ implies $G(\lambda^1) \geq G(\lambda^2)$ [23]. Equivalently, Schur-concave functions are characterized by being monotonic with respect to the majorization relation. They include all functions $G(\cdot)$ that are concave and symmetric with respect to the arguments, a subset which we have defined as quasi-entropic. An important subclass of quasi-entropic functions is obtained if we restrict G to have the following trace structure:

$$G(\lambda) = \text{tr}[g(\lambda)] = \sum_{\alpha} g(\lambda_{\alpha}), \quad (20)$$

where $g(\cdot)$ is a concave function of a single variable and λ is treated as a diagonal matrix for the purpose of calculating the trace. Note that $G(\cdot)$ as constructed in Eq. (20) is manifestly symmetric and, as a sum of concave functions, it is also concave. We shall refer to this class of functions, which include the Shannon and Tsallis entropies, as *trace type* quasi-entropic. An important fact regarding this class of measures is that if, for a given pair of

probability vectors λ and μ , we have $G(\lambda) > G(\mu)$ for every trace type quasi-entropic measure G , then $\lambda \prec \mu$.

The standard entropic measure of uncertainty [9], which is based on the Shannon entropy function, is trace type quasi-entropic and corresponds to the choice $g(x) = H(x) = -x \ln(x)$ in Eq. (20). An example of an information theoretically relevant measure that is quasi-entropic but not trace type is the Rényi subfamily of entropies of order less than one [24]. Accordingly, although the scalar entanglement detectors that follow from the theorems of §III actually hold for the entire class of Schur-concave measures, we will continue to focus on the quasi-entropic class as the most suitable for information theoretical applications.

Scalar entanglement detectors based on Shannon, Tsallis, and Rényi entropies are already well known, albeit with detection bounds that may differ from those reported here [5, 7, 8]. Our primary purpose here is to emphasize the natural and categorical manner in which majorization based entanglement detectors yield entire classes of scalar detectors.

With $G(\cdot)$ a quasi-entropic function, the aforementioned monotonicity property implies the following corollaries to Theorems 1-3.

Corollary 1. Under the conditions of Theorem 1, the multipartite state $\hat{\sigma}^{ABC\dots F}$ is entangled if

$$G[\mathcal{P}^X(\hat{\sigma}^{ABC\dots F})] < G[\mathcal{P}_{sep;sup}^X]. \quad (21)$$

Corollary 2. Under the conditions of Theorem 2, the bipartite state $\hat{\sigma}^{AB}$ is entangled if

$$G[\mathcal{P}^{(X^A \otimes X^B)}(\hat{\sigma}^{AB})] + G[\mathcal{P}^{(Y^A \otimes Y^B)}(\hat{\sigma}^{AB})] < G[\mathcal{P}_{sup}^{X \otimes Y}]. \quad (22)$$

Corollary 3. Under the conditions of Theorem 3, the multipartite state $\hat{\sigma}^{ABC\dots F}$ is entangled if

$$G[\lambda^{ABC\dots F}] < G[\Lambda^{ABC\dots F}]. \quad (23)$$

As an application of the above corollaries, we will consider the entanglement detection threshold for the bipartite Werner state considered in Eq. (8) et seq. Using Corollary 1 together with the Tsallis entropy function, we find that $\hat{\rho}_d^{wer}(q)$ is entangled if

$$S_r^{tsa}[\mathcal{P}^{X^*}[\hat{\rho}_d^{wer}(q)]] < S_r^{tsa}[\mathcal{P}_{sep;sup}^{X^*}], \quad (24)$$

where $S_r^{tsa}(\cdot)$ is the Tsallis entropy of order r , and the two probability vectors appearing in Eq. (24) are those in Eqs. (10) and (11). The Tsallis entropy function $S_r^{tsa}(\cdot)$ corresponds to the choice $g(\lambda) = (\lambda - \lambda^r)/(r-1)$, $1 < r < \infty$, in Eq. (20). The Tsallis entropy of order 1 is defined by continuity and equals the Shannon entropy [25].

A straightforward calculation turns Eq. (24) into

$$\begin{aligned} & \frac{1 - [q + (1 - q)/d^2]^r - (d^2 - 1)[(1 - q)/d^2]^r}{r - 1} \\ & < \frac{1 - d^{1-r}}{r - 1}. \end{aligned} \quad (25)$$

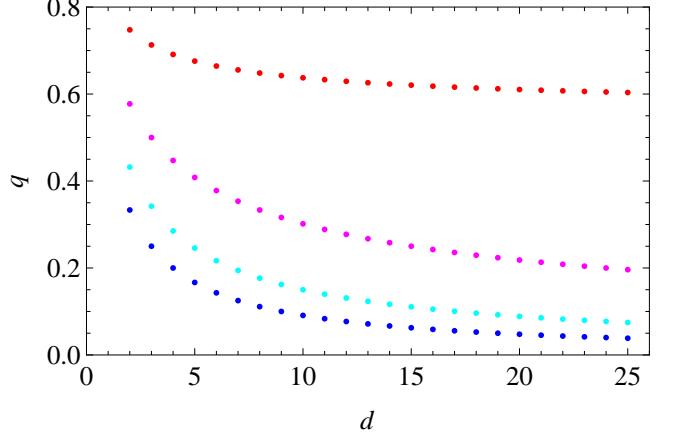


FIG. 1: Entanglement detection threshold q for a bipartite Werner state versus dimension d of each party, using Corollary 1 and the Tsallis entropy. The order of the Tsallis entropy equals, from top to bottom respectively, 1(red), 2 (magenta), 5 (turquoise), and ∞ (blue). The top graph corresponds to the Shannon entropy and the bottom one (d) reproduces the known entanglement threshold for the state.

For each value of r , Eq. (25) yields an entanglement detection threshold value for q which depends on d , and for a fixed value of the latter, decreases with increasing r . This corresponds to improving detection performance with increasing r , a behavior which is known for d equal to 2 and 3 [8]. Figure 1 shows a plot of the threshold value q versus the dimension d for four different values of the order r , respectively increasing from top to bottom. The improvement in entanglement detection for fixed d and increasing r , as well as fixed r and increasing d , is clearly in evidence in Fig. 1.

The weakest performance obtains for $r = 1$, top graph in Fig. 1, and can be found by taking the limit of Eq. (25) as $r \rightarrow 1$. The result is $H[q + (1 - q)/d^2] + (d^2 - 1)H[q + (1 - q)/d^2] < \ln(d)$, which corresponds to the choice of Shannon entropy function $H(\cdot)$ together with the set of generalized Bell states for the measurement of $d \otimes d$ Werner states [26]. The best performance obtains for $r \rightarrow \infty$ in Eq. (25), which simplifies to $q > 1/(1 + d)$, the known entanglement threshold for a the two-qudit Werner state. This limiting case corresponds to the bottom graph in Fig. 1.

The foregoing discussion of how well Corollary 1 does in detecting entanglement notwithstanding, the aim of this section is not to promote quasi-entropic detectors given in Corollaries 1-3 per se, but rather to emphasize the reach of the majorization results of §III whence they are inherited. In addition, the above example highlights the fact that, while the entanglement detection bounds given in Corollaries 1-3 are optimal for Schur-concave functions as a class, their performance need not be optimal for individual members of the class, a fact that was emphasized in paper I as well.

V. CONCLUDING REMARKS

In this paper we have developed entanglement detection criteria within the majorization formulation of uncertainty presented in paper I. These are inherently stronger than similar scalar conditions in the sense that they are equivalent to and imply infinite classes of such scalar criteria. Majorization relations in effect envelope the set of scalar measures of disorder based on Schur-concave functions, a huge set that includes the information-theoretically relevant class of quasi-entropic measures as a subset. This enveloping property elucidates the exceptional effectiveness of majorization relations in dealing with problems of quantum information theory.

Entanglement detection criteria that can be experimentally implemented are especially useful in studies involving entangled microsystems, hence the importance of linear detectors and entanglement witnesses. As already mentioned, Theorem 1 can be formulated as an entanglement witness inasmuch as it is a linear majorization condition on measurement results. What may not be so obvious is that in principle the spectrum of a state is also experimentally accessible on the basis of the Lemma of §IIIA. That Lemma states that the spectrum of a quantum state is the supremum of probability vectors that may be deduced from the counter statistics of all possible rank 1 generalized measurements performed on the state. In practice, this provides an experimental method for estimating the spectrum, since a high precision determination may require a search over a large number of possible measurements.

There is already an extensive literature on entanglement detection with many important results. With the notable exception of the Nielsen-Kempe theorem, these results are not formulated within the majorization framework. The present contribution in part serves to demonstrate the effectiveness of the majorization formulation of some of the existing methods. One can reasonably expect a sharpening of the results of some of the other entanglement detection strategies when reformulated within the majorization framework. It is also not unreasonable to expect useful, albeit computationally intractable, necessary and sufficient separability criteria to emerge from such formulations.

Acknowledgments

This work was in part supported by a grant from California State University, Sacramento.

Appendix: majorization bounds are found on pure states

Our task here is to establish that majorization uncertainty bounds can be reached on pure states. We will

prove this for two important cases, first for the uncertainty bound on all states of the system, and second on the class of separable states. To avoid unnecessary clutter and technical distractions, we will limit the number of generalized measurements to three and the underlying Hilbert spaces to finite dimensions in the following treatment.

Suppose generalized measurements $(X, \{\hat{E}_\alpha^X\})$, $(Y, \{\hat{E}_\beta^Y\})$, and $(Z, \{\hat{E}_\gamma^Z\})$ are performed on a state $\hat{\rho}$ that is supported on a finite-dimensional Hilbert space. What we will show below is that the search for the components of μ_i^{sup} which serve to define $\mathcal{P}_{sup}^{X \oplus Y \oplus Z}$ can be limited to pure states.

By definition,

$$\mathcal{P}_{sup}^{X \oplus Y \oplus Z} = \sup_{\hat{\rho}} [\mathcal{P}^{X \oplus Y \oplus Z}(\hat{\rho})], \quad (A.1)$$

where

$$\mathcal{P}_{\alpha\beta\gamma}^{X \oplus Y \oplus Z}(\hat{\rho}) = \text{tr}(\hat{E}_\alpha^X \hat{\rho}) \text{tr}(\hat{E}_\beta^Y \hat{\rho}) \text{tr}(\hat{E}_\gamma^Z \hat{\rho}). \quad (A.2)$$

We recall from §IIA above and §IIB, IVA and IVB of Paper I that the supremum defined by Eqs. (A.1) and (A.2) is found by calculating μ_i^{sup} for $i = 1, 2, \dots$ ($\mu_0^{sup} = 0$ by definition). Furthermore, μ_1^{sup} is the maximum value of a single component of $\mathcal{P}^{X \oplus Y \oplus Z}(\hat{\rho})$ as $\hat{\rho}$ is varied, μ_2^{sup} is the maximum value of the sum of two (different) components of $\mathcal{P}^{X \oplus Y \oplus Z}(\hat{\rho})$ as $\hat{\rho}$ is varied, and μ_i^{sup} is the maximum value of the sum of i (different) components of $\mathcal{P}^{X \oplus Y \oplus Z}(\hat{\rho})$ as $\hat{\rho}$ is varied. It is thus sufficient to prove that the said maximum for μ_i^{sup} is in fact realized on a pure state.

Let $\hat{\rho} = \sum_a \lambda_a(\hat{\rho}) |\psi_a\rangle\langle\psi_a|$ be the principal ensemble representation for $\hat{\rho}$. Then the above maximization process defines μ_i^{sup} as the maximum of

$$\begin{aligned} & \sum_{a,b,c} \lambda_a(\hat{\rho}) \lambda_b(\hat{\rho}) \lambda_c(\hat{\rho}) [\langle\psi_a | \hat{E}_{\alpha_1}^X | \psi_a\rangle \langle\psi_b | \hat{E}_{\beta_1}^Y | \psi_b\rangle \\ & \times \langle\psi_c | \hat{E}_{\gamma_1}^Z | \psi_c\rangle + \langle\psi_a | \hat{E}_{\alpha_2}^X | \psi_a\rangle \langle\psi_b | \hat{E}_{\beta_2}^Y | \psi_b\rangle \\ & \times \langle\psi_c | \hat{E}_{\gamma_2}^Z | \psi_c\rangle + \dots \\ & + \langle\psi_a | \hat{E}_{\alpha_i}^X | \psi_a\rangle \langle\psi_b | \hat{E}_{\beta_i}^Y | \psi_b\rangle \langle\psi_c | \hat{E}_{\gamma_i}^Z | \psi_c\rangle] \\ & + \xi_i (1 - \sum_a \lambda_a^\rho) + \sum_{a,b} \eta_{i,ab} (\delta_{ab} - \langle\psi_b | \psi_a\rangle), \end{aligned} \quad (A.3)$$

as the state vectors $\{|\psi_a\rangle\}$ which are the eigenstates of $\hat{\rho}$, the probabilities $\{\lambda_a(\hat{\rho})\}$ which constitute the spectrum of $\hat{\rho}$, and the Lagrange multipliers ξ_i and $\eta_{i,ba}$ which serve to enforce the structure of $\sum_a \lambda_a^\rho |\psi_a\rangle\langle\psi_a|$ as an orthogonal ensemble are varied. The symbol δ_{ab} in the above expression is the Kronecker delta, and the three sets of indices $\{(\alpha, \beta, \gamma)_1, (\alpha, \beta, \gamma)_2, \dots, (\alpha, \beta, \gamma)_i\}$ are understood to be distinct, i.e., $\alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_i$, and similarly for the other two sets. Note also that since $\langle\psi_a | \psi_b\rangle$ is in general a Hermitian matrix (in the indices a and b), the Lagrange multipliers $\eta_{i,ba}$ may also be taken to constitute a Hermitian matrix.

A variation with respect to $\langle\psi_a |$ gives

$$\lambda_a(\hat{\rho}) [\hat{E}_i^X + \hat{E}_i^Y + \hat{E}_i^Z] |\psi_a\rangle = \sum_b \eta_{i,ba} |\psi_b\rangle, \quad (A.4)$$

where

$$\hat{\mathcal{E}}_i^X = \sum_{k=1}^i \mathcal{P}_{\beta_k}^Y(\hat{\rho}) \mathcal{P}_{\gamma_k}^Z(\hat{\rho}) \hat{E}_{\alpha_k}^X, \quad (\text{A.5})$$

and $\hat{\mathcal{E}}_i^Z$ and $\hat{\mathcal{E}}_i^Y$ are defined similarly. Note that the three operators introduced in Eqs. (A.4) and (A.5) are Hermitian and positive.

Let $\hat{\mathcal{E}}_i = \hat{\mathcal{E}}_i^X + \hat{\mathcal{E}}_i^Y + \hat{\mathcal{E}}_i^Z$ and $\mathcal{E}_{i,aa} = \langle \psi_a | \hat{\mathcal{E}}_i | \psi_a \rangle$. Then the maximum sought in (A.3), namely μ_i^{sup} , is given by $\sum_a \lambda_a(\hat{\rho}) \mathcal{E}_{i,aa} / 3 = \text{tr}[\hat{\mathcal{E}}_i \hat{\rho}] / 3$. A key fact at this juncture is that the diagonal elements $\mathcal{E}_{i,aa}$ do not depend on a and are all equal. In other words, each pure state $|\psi_a\rangle$ in the ensemble contributes equally to μ_i^{sup} . This property follows from a variation of (A.3) with respect to $\lambda_a(\hat{\rho})$, which leads to

$$\xi_i = \langle \psi_a | \hat{\mathcal{E}}_i^X + \hat{\mathcal{E}}_i^Y + \hat{\mathcal{E}}_i^Z | \psi_a \rangle = \mathcal{E}_{i,aa}. \quad (\text{A.6})$$

Clearly then, each pure state $|\psi_a\rangle$ in the ensemble must realize the same maximum μ_i^{sup} as the entire ensemble $\hat{\rho}$. We conclude therefore that the sought maximum can be found among the pure states of the system.

What if we are looking for μ_i^{sup} but with $\hat{\rho}$ limited to separable states? It turns out that here too the search can be limited to pure states, which would be pure product states in this instance. To establish this result, we will modify the foregoing analysis by stipulating that $\hat{\rho} = \sum_a q_a |\phi_a^A\rangle\langle\phi_a^A| \otimes |\phi_a^B\rangle\langle\phi_a^B|$ representing a separable, bipartite state of parties A and B . Then the corresponding $\mu_{sep,i}^{sup}$ is the maximum of

$$\begin{aligned} & \sum_{a,b,c} q_a q_b q_c [\langle \phi_a^B | \otimes \langle \phi_a^A | \hat{E}_{\alpha_1}^X | \phi_a^A \rangle \otimes | \phi_b^B \rangle \\ & \quad \times \langle \phi_b^B | \otimes \langle \phi_b^A | \hat{E}_{\beta_1}^Y | \phi_b^A \rangle \otimes | \phi_b^B \rangle \\ & \quad \times \langle \phi_c^B | \otimes \langle \phi_c^A | \hat{E}_{\gamma_1}^Z | \phi_c^A \rangle \otimes | \phi_c^B \rangle \\ & \quad + \langle \phi_a^B | \otimes \langle \phi_a^A | \hat{E}_{\alpha_2}^X | \phi_a^A \rangle \otimes | \phi_a^B \rangle \\ & \quad \times \langle \phi_b^B | \otimes \langle \phi_b^A | \hat{E}_{\beta_2}^Y | \phi_b^A \rangle \otimes | \phi_b^B \rangle \\ & \quad \times \langle \phi_c^B | \otimes \langle \phi_c^A | \hat{E}_{\gamma_2}^Z | \phi_c^A \rangle \otimes | \phi_c^B \rangle \\ & \quad + \dots \\ & \quad + \langle \phi_a^B | \otimes \langle \phi_a^A | \hat{E}_{\alpha_i}^X | \phi_a^A \rangle \otimes | \phi_a^B \rangle \\ & \quad \times \langle \phi_b^B | \otimes \langle \phi_b^A | \hat{E}_{\beta_i}^Y | \phi_b^A \rangle \otimes | \phi_b^B \rangle \\ & \quad \times \langle \phi_c^B | \otimes \langle \phi_c^A | \hat{E}_{\gamma_i}^Z | \phi_c^A \rangle \otimes | \phi_c^B \rangle \\ & + \sum_a \eta_{i,a}^A [1 - \langle \phi_a^A | \phi_a^A \rangle] + \sum_a \eta_{i,a}^B [1 - \langle \phi_a^B | \phi_a^B \rangle] \\ & \quad + \xi_i (1 - \sum_a q_a), \end{aligned} \quad (\text{A.7})$$

as the state vectors $\{|\phi_a^A\rangle\}$, $\{|\phi_a^B\rangle\}$, the probabilities $\{q_a\}$, and the Lagrange multipliers $\eta_{i,a}^A$, $\eta_{i,a}^B$, and ξ_i which serve to enforce normalization conditions and probability conservation for the ensemble are varied.

Variations with respect to $\langle \phi_a^A |$ and $\langle \phi_a^B |$ now give

$$\begin{aligned} q_a [\langle \phi_a^B | (\hat{\mathcal{E}}_i^X + \hat{\mathcal{E}}_i^Y + \hat{\mathcal{E}}_i^Z) | \phi_a^B \rangle] | \phi_a^A \rangle &= \eta_{i,a}^A | \phi_a^A \rangle \\ q_a [\langle \phi_a^A | (\hat{\mathcal{E}}_i^X + \hat{\mathcal{E}}_i^Y + \hat{\mathcal{E}}_i^Z) | \phi_a^A \rangle] | \phi_a^B \rangle &= \eta_{i,a}^B | \phi_a^B \rangle, \end{aligned} \quad (\text{A.8})$$

where $(\hat{\mathcal{E}}_i^X, \hat{\mathcal{E}}_i^Y, \hat{\mathcal{E}}_i^Z)$ are defined as in Eq. (A.5) et seq.

Equations (A.8) directly imply that $\eta_{i,a}^A = \eta_{i,a}^B$, and as in the general case above, the maximum sought in (A.7), namely $\mu_{sep,i}^{sup}$, is found to equal $\sum_a q_a \mathcal{E}_{i,aa} / 3 = \text{tr}[\hat{\mathcal{E}}_i \hat{\rho}] / 3$, where $\hat{\mathcal{E}}_i = \hat{\mathcal{E}}_i^X + \hat{\mathcal{E}}_i^Y + \hat{\mathcal{E}}_i^Z$. Furthermore, a variation with respect to the probabilities $\{q_a\}$ shows the equality of the matrix elements $\mathcal{E}_{i,aa}$ just as in the general case above, whereupon we learn that each pure product state in the ensemble makes the same contribution to $\mu_{sep,i}^{sup}$. In other words, the separable density matrix which maximizes (A.7) may be taken to be a pure product state [16].

It should be clear from the above arguments that the results hold for any number of generalized measurements as well as any number of parties in the multipartite state. In summary, then, we have found that the joint majorization uncertainty bound, $\mathcal{P}_{sup}^{X \oplus Y \oplus \dots \oplus Z}$, for a set of generalized measurements performed on an arbitrary quantum state supported on a finite-dimensional Hilbert space can be found on the class of pure states of the system, and in the case of a separable states, $\mathcal{P}_{sep;sup}^{X \oplus Y \oplus \dots \oplus Z}$ can be found on the class of pure product states of the system.

[1] W. Heisenberg, *Zeit. Phys.* **43**, 172 (1927).
[2] See S. Wehner and A. Winter, *New J. Phys.* **12**, 025009 (2010), for a survey and references to the literature.
[3] M. Hossein Partovi, *Phys. Rev. A* **84**, 052117 (2011).

The literature on the subject can also be traced from the references given in this paper.
[4] L. Gurvits, *Proceedings of the Thirty-fifth ACM Symposium on the Theory of Computing* (ACM Press, New

York, 2003), pp. 10-19.

[5] The extensive literature on entanglement detection can be traced from the review article by O. Guhne and G. Tth, *Physics Reports* **474**, 1 (2009).

[6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).

[7] V. Giovannetti, *Phys. Rev. A* **70**, 012102 (2004).

[8] Guhne and M. Lewenstein, *Phys. Rev. A* **70**, 022316 (2004).

[9] D. Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983).

[10] A good reference to majorization theory is A. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications* (Academic Press, New York, 1979). For physics applications, see Michael Nielsen, *Phys. Rev. Lett.* **83**, 436 (1999); Michael Nielsen and J. Kempe, *Phys. Rev. Lett.* **86**, 5184 (2001); Michael Nielsen and G. Vidal, *Quant. Inf. Comp.* **1**, 76 (2001); M. Hossein Partovi, *Phys. Rev. Lett.* **103**, 230502 (2009), and *Phys. Rev. A* **84**, 052117 (2011).

[11] Because the complete Rnyi family fails to be concave, we classified it as not belonging to the quasi-entropic class in Paper I. Here we will use a more nuanced classification; see §IV for further discussion, as well Ref. [24] below.

[12] M. Hossein Partovi, *Phys. Rev. Lett.* **103**, 230502 (2009).

[13] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[14] See the paper by Nielsen and Kempe in Ref. [10].

[15] Note that by an “entangled state” we mean any multipartite state that is not fully separable.

[16] In technical language, these pure product states are extremal points in the (convex) set of separable states.

[17] M. Lewenstein, B. Kraus, J. Cirac, and P. Horodecki, *Phys. Rev. A* **62**, 052310 (2000).

[18] R. Werner, *Phys. Rev. A* **40**, 4277 (1989).

[19] A. Pittenger and M. Rubin, *Optics Commun.* **179**, 447 (2000); P. Rungta, W.J. Munro, K. Nemoto, P. Deuar, G.J. Milburn, and C.M. Caves, in *Directions in Quantum Optics: A Collection of Papers Dedicated to the Memory of Dan Walls*, edited by H.J. Carmichael, R.J. Glauber, and M.O. Scully (Springer, Berlin, 2001), p. 149.

[20] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Guhne, P. Hyllus, D. Bru, M. Lewenstein, and A. Sanpera, *Phys. Rev. Lett.* **92**, 087902 (2004), the Appendix.

[21] The ensuing condition here is precisely equivalent to the separability condition in Eq. (8) of Nielsen and Kempe [10] which was obtained by a different method (to be discussed in §IIIB). Needless to say, this is not a surprising outcome since both methods are based on majorization relations.

[22] Here we are using the readily proven theorem that $\lambda_1 \prec \mu_1$ and $\lambda_2 \prec \mu_2$ imply $\lambda_1 \otimes \lambda_2 \prec \mu_1 \otimes \mu_2$. Note that the converse to this theorem is false, a fact that makes entanglement catalysis possible; see D. Jonathan and M. Plenio, *Phys. Rev. Lett.* **83**, 3566 (1999).

[23] See the book cited under Ref. [10].

[24] The Tsallis family is of the trace type quasi-entropic class defined in Eq. (20). The Rnyi family, on the other hand, is not of the trace type and in addition has mixed concavity properties: it is Schur concave and quasi-entropic if its order is no greater than 1 (and equals the Shannon entropy for order 1), but fails to be concave or Schur concave if its order exceeds 1; see, e.g., P. A. Varotsos, N. V. Sarlis, H. K. Tanaka, and E. S. Skordas, *Phys. Rev. A* **71**, 032102 (2005).

[25] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).

[26] Similar results have been found by other methods; see [7].